DOCUMENT RESUME

ED 186 746 CE 025 403

AUTHOR	McCray, Paul
TITLE	An Interpretation of VEWAA/CARF Work Sample Standards.
INSTITUTION	Wisconsin Univ. - Stout, Menomonie. Dept. of Rehabilitation and Manpower Services. Materials Development Center.
SPONS AGENCY	Rehabilitation Services Administration (DHEW), Washington, D.C.
PUB DATE	Jul 79
GRANT	DHEW-RSA-12-P-55307/5
NOTE	24p.; For a related document see CE 025 411.
AVAILABLE FROM	Materials Development Center, Stout Vocational Rehabilitation Institute, University of Wisconsin-Stout, Menomonie, WI 54751 ($1.00).
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Diagnostic Tests; Job Performance; Job Skills; Occupational Tests; *Situational Tests; *Standards; Test Validity; Vocational Adjustment; *Vocational Aptitude; *Vocational Rehabilitation; *Work Sample Tests
IDENTIFIERS	*Vocational Evaluation

ABSTRACT

The purpose of this publication is to review each of the Vocational Evaluation and Work Adjustment Association sample testing standards as well as to provide insight and understanding into the underlying purposes of the guidelines. The first of three sections focuses on those standards designed to insure that a work sample actually measures what it purports to measure and is representative of real competitive worker skills. Section 2 describes the requirements which define what constitutes work sample standardization. Specific requirements discussed are job analysis relationship; prerequisites; work sample purpose; materials and equipment used; test preparation and layout; instructions; time, error, and scoring instructions; and instructions for interpreting scores. The final section expounds upon the standard which requires work samples to have appropriate competitive or industrial norms. (LRA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED186746

AN INTERPRETATION OF VEWAA/CARF WORK SAMPLE STANDARDS

By

Paul McCray, M.S.

CE025403

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. Fry

TO THE EDUCATIONAL RESOURCES NFORMATION CENTER (ERIC)."

ADDITIONAL COPIES OF THIS PUBLICATION
MAY BE PURCHASED FROM MDC FOR $1.00 PER COPY

AN INTERPRETATION OF VEWAA/CARF WORK SAMPLE STANDARDS

Introduction

Of the four major vocational evaluation techniques--work samples, psychological testing, situational assessment, and job site evaluation--work sample testing has become one of the most often used methods. Indeed, in many respects it is a process that distinguishes the field of vocational evaluation from other more traditional psychological testing approaches. Yet with the increased application of this assessment technique, a corresponding responsibility to assure that it is carried on in a professional manner has arisen. In an effort to meet this charge, the Vocational Evaluation and Work Adjustment Association (VEWAA) has developed minimum standards to be applied to work samples. These standards have been adopted by the Commission on Accreditation of Rehabilitation Facilities (CARF) because they provide a sound basis for the development of reliable work samples.

The present CARF work sample guidelines read as follows:

3.4.3.1.1.7.2 If work samples are used:

a. the vocational evaluation service work samples resources shall be representative of realistic competitive worker skills.

b. work samples shall be established by an analysis of job tasks or traits related to a specific area of work, and be standardized as to materials, layout, instructions, and scoring.

c. competitive norms or industrial standards shall be established and used. (p. 28)

Thus, it is clear that the guidelines for some of the content, the structure, and the development of work samples are relatively specific. The purpose of this publication is to review each of the work sample testing standards as well as provide insight and understanding into the underlying purposes of the guidelines. With this information in mind, vocational evaluators and program administrators should be able to better assess the effectiveness and quality of their work samples with regard to: (1) meeting the needs of their clients and referral sources in terms of using valid and reliable assessment techniques, and (2) satisfying professional standards as well as CARF accreditation standards for vocational evaluation programs. In addition, the guidelines should serve as a valuable reference for selecting work samples which satisfy these requirements and, therefore, are most likely to provide reliable and valid information as to client capabilities and limitations.

Paul McCray, M.S.
July, 1979

Table of Contents

	Page
PART I	
Work Sample Representativeness	1
PART II	
Work Sample Standardization	6
PART III	
Work Sample Norms	11
Norms and Industrial Standards	12
Work Measurement Applied to Work Samples	15
Cautions in Using Competitive Norms or Industrial Standards	17
SUMMARY	19
REFERENCES	20

PART I

Work Sample Representativeness

Standard 3.4.3.1.1.7.2a of the CARF Standards Manual for Rehabilitation Facilities indicates that if work samples are used, "the vocational evaluation service work samples resources shall be representative of realistic competitive worker skills" (p. 28). A joint CARF-VEWAA interpretation of this standard goes on to explain:

> Evaluation programs using work samples that purport to be representative of specific work skills should be able to provide documentation that they do. Documentation might include the considered judgement of appropriate employers or employees as to the ability of a work sample to measure skills related to a particular occupation. (Vocational Evaluation and Work Adjustment Association and Commission on Accreditation of Rehabilitation Facilities, 1978, p. 6)

The major implication of this standard and the interpretation is that work samples should have evidence of validity. There must be documentation that a given work sample really measures what it purports to measure. In other words, if a work sample says that it measures a client's ability to work as a mail sorter, there should be evidence that it actually does so. In essence, this requires that the work sample contain the critical elements necessary for satisfactory mail sorting performance, and that they can be reliably measured and analyzed. Although knowledge of how to establish a test's validity is certainly important, it is not the purpose of this publication to discuss the various measures of validity and how they are related to work samples. Interested readers are referred to Dunn (1971) for a brief review of this subject.

The phrase "representative of realistic competitive worker skills" is of critical importance to understanding this guideline. It refers to the idea that a work sample should measure the critical elements, e.g., aptitudes, traits, and skills, actually required to perform a job found in the labor market. It should not measure elements that are not required for satisfactory performance, rather it should include elements which are necessary for competitive performance. For example, if a work sample contains all the elements required to perform the job of lathe operator, one could certainly say that it represents real competitive worker skills. On the other hand, if the same work sample required the client to perform a task which was not part of the actual job, e.g., memorize an operator's manual, one should recognize that the work sample is no longer representative of real work requirements, because the ability to memorize a manual is not actually required on the job. Similarly, if a critical aspect of the job is the ability to read blueprints, and this skill is not included in the work sample, then it is also not fully representative of all the real requirements of the job.

Therefore, if a work sample purports to measure one's ability to work in a given job or job area, it must contain the crucial elements required for satisfactory job performance. Likewise, if a work sample claims to measure a given trait, that trait should be clearly defined and there should be evidence that the trait has been accurately incorporated in the work sample and is systematically measured. If the individual traits are described as being

critical for satisfactory job performance, there must be documentation that it is actually true. Such evidence may be obtained from job analysis, task analysis, Dictionary of Occupational Titles (DOT) data, etc.

The standard also indicates that a work sample should be realistic. Basically this means that it should relate to actual work. Thus, when a client is performing a work sample, he can see that the work he is doing is similar to competitive work which really exists in industry. This lends credibility to the assessment instrument in that it is no longer an abstract concept or process that lacks meaningfulness for the client. This is one of the advantages work samples have over paper and pencil tests which clients often mistakenly view as lacking any relationship to real work.

Once it is recognized that work samples should contain the elements that are crucial to satisfactory job performance, as well as be realistic, it follows that one must ask how these elements can be objectively identified. The first half of part "b" of the standard points out the method for identifying realistic competitive worker skills. It reads: "work samples shall be established by an analysis of job tasks or traits related to a specific area of work" (p. 28). Job analysis is probably the most effective and efficient way to determine the critical aspects of a job. It essentially consists of analyzing the specific job related tasks, aptitudes, physical demands, machines, tools, equipment, interests, etc., required for satisfactory job performance. Once this information is available, the evaluator may begin to incorporate these elements into the work sample. Certainly the evaluator should attempt to include as many of the job tasks in the work sample as is practically feasible. If the work sample entirely replicates a job performed in industry in terms of equipment, tools, production methods, work standards, etc., it is better thought of as a job sample. The VEWAA-CARF glossary defines job sample as:

> Job Sample - Those worksamples that in their entirety are replicated directly from industry and include the equipment, tools, raw materials, exact procedures, and work standards of the job. (Vocational Evaluation and Work Adjustment Association and Commission on Accreditation of Rehabilitation Facilities, 1978, p. 20)

A job sample can be very useful in assessing a client's potential to do a specific job, but it still has limitations; the primary one being that the work environment may not be entirely duplicated. This is a difficult and often impossible task, since there is such a large number of uncontrolled variables that make up any given work setting. But to the maximum extent possible, work and job samples should simulate the work environment along with the work tasks as closely as possible. In any case, it is imperative that the critical job elements be included in the work sample. Otherwise, a client may be able to perform well on a work sample, but cannot actually perform the job. For example, although a bricklayer may only spend a small portion of his time measuring a building layout, this skill must be included in the work sample, because without it the client will not be able to work competitively. Thus, the term critical refers to those elements of a job which are essential to satisfactory overall job performance. An example of a noncritical element which need not be incorporated in a bricklayer work sample is fine finger dexterity. Although a bricklayer might have to periodically manipulate small objects, it generally need not be done at exceptionally rapid rates such as are common to assembly type operations. The evaluator is primarily concerned

with making sure that the client simply can manipulate small objects. A client's ability to do such a task can generally be intuitively assessed by an evaluator.

Since it is often unrealistic to expect a work sample to contain all the elements of a job, it is important that somewhere in the manual the elements which were not included in the work sample be listed. In this way the evaluator is made aware of the limitations of the work sample and in many cases, he will be able to utilize different combinations of work samples to supplement each other. For example, if a job such as photocopy machine operator requires the worker to operate a cash register 5% of the time, it may not be feasible to include an expensive cash register as part of the work sample. Yet once this problem is pointed out, many evaluation programs which already possess cashiering assessment work samples will be able to supplement the photocopy work sample in such a way that a more reliable and thorough assessment of the client's abilities is possible.

Work samples should not only have the critical elements of a job, but they should also contain those elements in a proportion similar to that of the job. Thus, if a telephone solicitor spends 90% of his time talking on the phone and 10% of the remaining time recording information, the work sample should reflect the proportional relationship of the elements. Therefore, a telephone solicitor work sample should not have a client spend 30% of his time on the phone and 70% of the time recording information. Instead, the 90/10% relationship would be incorporated into the work sample so as to provide a more realistic simulation of the job.

Other techniques besides job analysis are also available for determining job elements. Task analysis is one method. It basically involves analyzing in a step-by-step manner the procedures involved in the individual tasks that make up a job. In other cases, evaluators may be able to obtain detailed job descriptions which specify the traits, aptitudes, etc., required as well as the amount of time employees are expected to spend on specific job tasks for satisfactory performance. This information may be available from the personnel department of public and private businesses, state employment service offices, or occupational analysis field laboratories of the Department of Labor. Another method is to analyze the job descriptions, worker traits, worker functions, etc., provided in the DOT. However, the major weakness with this approach is that the DOT information may differ from the individual job requirements of local industry. In any case, one of the most important reasons for analyzing a job should be to accurately determine the elements that are important to a job so that they may be incorporated in a work sample. As mentioned previously, this process not only helps insure the representativeness of realistic competitive worker skills, but also provides documented evidence that an analysis of job requirements has taken place.

Although the importance of insuring representativeness as well as how this goal can be achieved has been pointed out, there has been no discussion as to the problem of how many or how much of the actual job elements must be included in a work sample. For example, once an analysis of the job has been completed and as many of the elements as possible have been incorporated into the work sample, the evaluator can compare the degree to which the work sample contains the elements of the job. Look at the following example:

Telephone Solicitor, DOT Code 299.357-014

Job Analysis Description of Tasks	Work Sample Tasks
Calls prospective customers on phone, explains merchandise, and solicits purchases. (70%)	Calls mock customers on phone, explains merchandise and solicits purchases. (70%)
Records names and addresses of purchasers as well as amount of purchase. (15%)	Records names and addresses of purchasers as well as amount of purchase. (15%)
Reads telephone book to obtain phone numbers of prospective customers. (10%)	The other 15% of the work sample involves non-related activities primarily consisting of cleaning the work area and replacing materials.
Files orders. (5%)	

One can see that the work sample contains approximately 85% of the job elements which were derived from a job analysis on the job of telephone solicitor. The obvious question that arises is what percentage of work sample/job overlap is adequate for a work sample to be considered representative of realistic worker skills? Certainly, it is almost impossible to incorporate all of the elements of almost any job into work samples, since most jobs contain a myriad of task-related and behaviorally-related factors which can only exist in the real work environment. Thus, it will be very rare for there to be a 100% match-up between job and work sample elements.

At the present time, there are no clear-cut guidelines as to what constitutes an adequate match-up. Indeed, all that has really been determined is that all the elements of a job need not be included in a test or work sample in order for the test to be representative or valid. Manning's (1978) interpretation of a recent Supreme Court decision indicates that, "'key elements,' 'critical skills,' or 'minimal competencies,' or 'significant components' of a job can be used in test validation instead of the full range of skills, knowledge, and other attributes revealed by formal job analysis" (p. 76). Thus, not all the factors involved in a job must be included in a work sample in order for it to be reasonably valid. However, it is recommended that to the maximum extent possible, all critical elements of a job should be included in a work sample if the tool is to be used as a basis for determining employment potential on a given job. If it is impossible to incorporate some of the critical job elements into the work sample, then perhaps other evaluation techniques should be used, i.e., job site evaluation or supplemental work samples.

Insuring that work samples are representative of realistic competitive worker skills is important for many reasons. First, if a work sample does not actually measure what it purports to measure, then decisions and interpretations as to employability are based on a non-valid assessment technique which is potentially detrimental to a client's rehabilitation progress. For example, if a work sample fails to include some critical job tasks, then a client may be able to perform the work sample, but when placed on the job he finds that there are some important job demands which he cannot satisfy. This

situation has a negative impact on the client, the job placement person, and the potential employer. It also reflects poorly on the vocational evaluation service. Conversely, if representativeness is not achieved, a client may be screened out of a job he can actually do simply because he performed poorly on a work sample which required additional skills that were irrelevant to job performance. For example, suppose a client is taking a soldering work sample which requires him to read a complicated electronic circuitry blueprint even though the soldering jobs available in that geographic area do not require blueprint reading skills. Workers only have to solder together two wires in a standardized, routine way. In this case, the addition of a critical skill that is not a job requirement limits the representativeness of the work sample. This could easily lead to assuming that poor work sample performance on the blueprint reading indicates limited potential for employment as a solderer. Thus, the work sample is not representative of realistic local job requirements.

Finally, one of the most basic functions of vocational evaluation is to assess client interest in a specific job or occupational area. Therefore, in order for a true picture of client interest to emerge, the client's experience with the work sample must be based on an accurate representation of the job. Otherwise, client interest is based on a superficial or incorrect understanding of the job. This can lead to unsatisfactory job/training placement. For example, if an evaluator wants to assess a client's interest in the area of small engine mechanics, the work sample should simulate the job of small engine mechanic as closely as possible. This means that the client should work in an area with noxious odors, greasy and dirty parts, etc. However, in many cases, a client is given a shiny clean engine to work on in a quiet formal testing atmosphere. This presents an inaccurate picture of the work environment, thus, if a client expresses an interest in this job, it may be based on an inaccurate understanding of the job. In such a case, the client's interests may change dramatically once he is placed on the job. Therefore, if the evaluation setting does not do a reasonably good job of approximating the real work environment, the evaluator should give serious consideration to substituting job site evaluation techniques for work samples since the former is usually a better model of the job environment.

In summary, the primary rationale for CARF Standards 3.4.3.1.1.7.2.a and the first part of "b" is that they help insure that a work sample actually measures what it purports to measure and is representative of real competitive worker skills. Thus, the data which is derived from work sample testing provides a meaningful and accurate picture of a client's vocational potential. Then client, evaluator, and referral source have a valid and reliable source of information for decision making.

PART II

Work Sample Standardization

Part "b" of standard 3.4.3.1.1.7.2 states:

> work samples shall be established by an analysis of job tasks or traits related to a specific area of work, and be standardized as to materials, layout, instructions, and scoring. (CARF, 1978, p. 28)

This guideline actually has two sections. The first indicates that work samples are to be based on an analysis of a job or group of jobs, while the second suggests that work samples should be standardized. Part I of this publication has already explained the importance of analyzing a job in order to insure work sample representativeness. Therefore, Part II will be confined to a discussion of the concept of work sample standardization.

CARF has provided specific examples of what is to be included in a standardized work sample.

> Each work sample shall have an examiner's manual which specifies: (1) its relationship to the DOT, Occupational Divisions, Worker Trait Groups, or some appropriate job analysis system; (2) prerequisites, i.e., any specific work sample task requirements which might make administration of the sample unfeasible for a given individual; (3) the work sample purpose, i.e., specifically what is the sample attempting to assess; (4) the materials and equipment used; (5) preparations for testing and the layout of materials; (6) instructions to the individual; (7) instructions for timing, evaluating errors, and scoring if applicable; (8) instructions for interpreting scores. (p. 78)

From this interpretation one can see that there are some relatively explicit requirements which define what constitutes work sample standardization. The following examples provide a further clarification and interpretation for this guideline and readers are referred to the Materials Development Center (MDC) publication Work Sample Manual Format for further information.

1. Job Analysis Relationship

The job analysis relationship should be based on a standardized job information system which provides organized data as to the jobs, traits, work activities, etc., directly related to the work sample. The information might be presented in the following manner.

> This work sample was based on an analysis of the job of Watch Repairer (clock and watch), Dictionary of Occupational Titles (DOT), Fourth Edition, Code 715.281-010. In the DOT classification system, a Data-People-Things (DPT) Code of 281 indicates that the job requires an "analyzing" relationship to Data, a "taking instructions-helping" relationship to People, and a "precision working" relationship to Things (pp. 1369-1371).

According to the Worker Trait Group Guide (1978) of the Appalachia Educational Laboratory (AEL) system, the job of watch repairer falls in the Career Area of Industrial, Code 06., and the Worker Trait Group (WTG) of Production Technology, WTG 06.01 (pp. 189-195). Work activities include:

1. Activities dealing with things and objects . . .
9. Activities involving process, methods, or machines
10. Activities involving working on or producing things. (p. 193)

In this example, two different job description systems have been used, the DOT and the AEL. Both of these systems provide concrete and specific information as to job related factors. Additional data such as physical demands, aptitudes, interests, related jobs, etc., could also be included in this section.

2. Prerequisites

MDC (1977) defined work sample prerequisites as "minimum performance requirements from previous tests or work samples that will need to be met before the work sample may be administered" (p. 4). Examples of prerequisites might include: reading levels, physical and medical factors, specific aptitudes, educational training, etc. All prerequisites should be listed, e.g., be able to lift and carry 50 pound boxes, and any precautions such as safety hazards.

3. Work Sample Purpose

It is important to clearly identify in the work sample manual what specific factors the work sample is designed to assess. These factors should be as clearly stated as possible. General statements such as the purpose of this work sample is to assess "production skills," "clerical skills," "potential for competitive employment," etc., are too vague in that they don't indicate the exact skills, aptitudes, traits, etc., being assessed. Look at the following example:

Assessment Description - The purpose of this work sample is to assess the following factors:

1. Ability to work competitively as a drill press operator (wood).
2. Ability to use rulers, micrometers, and calipers to measure wood stocks.
3. Ability to read blueprints as they relate to drill press operators.
4. Safety habits with regard to operating a drill press.
5. Interest in working on routine, repetitive machine (drill press) operating controlling tasks.
6. Ability to stand for periods ranging up to three consecutive hours.
7. Spatial, form perception, motor coordination, finger dexterity, and manual dexterity as defined and rated by the Dictionary of Occupational Titles, Third Edition, Volume II, page 653.

This type of assessment description gives the evaluator as well as the client a clear-cut understanding of the specific purposes of the work sample. If a work sample assesses specific traits or other factors, then they might be listed in a similar fashion, e.g., spatial aptitudes, manual dexterity, GED, Data-People-Things codes, etc.

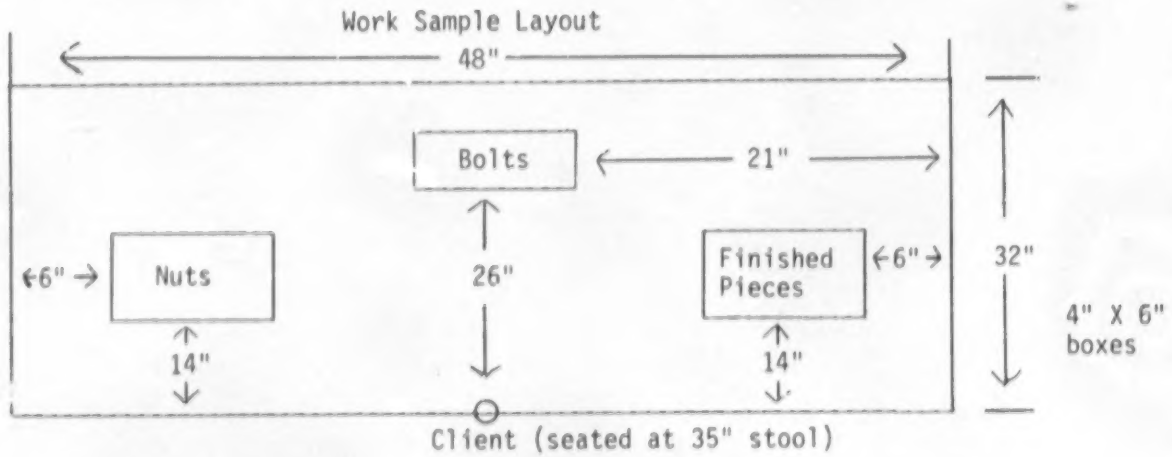
4. Materials and Equipment Used

All of the materials and equipment necessary in order to use a work sample should be listed in the manual. "Information including part names, quantities, purchase order numbers, shipping weights, sizes and descriptions such as color, grade, etc., purchase source and address, and total quantity price are necessary" (MDC, 1977, p. 26). This information helps insure that lost or worn-out parts and materials are replaced in a consistent manner so that the construction of the work sample does not change. If it should change, it may mean that norms, reliability, validity, job analysis information, etc., may no longer be applicable.

5. Test Preparation and Layout

Test preparation and layout basically refers to information related to work sample conditions, administration equipment, and setup and breakdown procedures.

The test environment should be described, e.g., indoors vs. outdoors, quiet vs. noisy, and each work sample should list the equipment that is necessary in order to proceed with the administration. Examples include: tools, hardware, safety devices, tables, manuals, scoring sheets, timers, etc. In addition to this, every work sample must be set up in a certain manner. The location of the tools, their arrangement, the method of setting up the work station, etc., should all be clearly spelled out. This will help to insure standardized administration procedures. If norms or industrial standards are used, it is particularly important to make sure that the work sample layout is the same as the layout used during the norming process. This means having objects positioned in well-defined locations with distances spelled out. In most cases, a diagram of the work sample layout is very helpful.



6. Instructions

Instructions to the individual refers to that information which is necessary for administration. It will often include some of the previously discussed items such as prerequisites, materials and equipment, and preparations for testing and layout. Instructions are generally provided for both the client and evaluator. The instructional procedures should be as specific as possible and should be standardized with regard to the format in which they are initially presented, e.g., written, oral, demonstration, and hands on.

If a client is unable to follow the standardized instructional format, it should be noted that other formats can be provided, but they need not be included in the work sample manual; for further discussion of this subject refer to McCray (1979). An example of instructions for both the client and evaluator are noted below:

> Client Instructions (to be read to the client by the evaluator) - "Your job is to take each of the bolts in this box (point) and thread a nut from this box (point) onto the bolt, and put the finished assembly into this box (point)."
>
> Evaluator Instructions - As the evaluator points to each of the first two boxes, he is to remove one sample object and show it to the client. At the same time he will assemble one nut and bolt and place it in the box labeled for finished assemblies.

Thus, it is evident that specific directions are to be provided; however, this example is by no means complete. If space permitted, it would go on to meticulously explain the exact production method, practice period, objectives, performance session requirements, etc.

7. Time, Error, and Scoring Instructions

This section refers to the basic scoring criteria which are used for analyzing a client's performance. The techniques to be used in scoring, quantity, quality, errors, etc., must be clearly identified along with all other measurable outcomes. In addition to this, information such as definitions as to what constitutes an error, when timing should begin and end, behavior observation requirements, definitions of relevant behaviors, etc., should also be included.

8. Instructions for Interpreting Scores

This information should provide special input into understanding a client's performance. It may include instructions as to how to read tables in order to identify appropriate norm groups. Other information might include how the norms were developed, whether they apply to experienced or non-experienced workers, learning curves, situational factors such as behavior problems, etc.

Work sample standardization is important for many reasons. First it helps insure that a client's performance is not unduly influenced by haphazard assessment procedures which may contribute to unreliable results. Second, it provides an objective basis for making consistent and valid observations because the skills, behaviors, concepts, etc., which the work sample is designed to assess are clearly defined and identified. Thus the evaluator and the client are both aware of the specific reason(s) for using the work sample. This adds meaningfulness to the technique and helps reduce the likelihood of individuals misunderstanding the purpose(s) of the test. Last, a standardized manual can be very useful for training new staff members. With a manual, should an evaluator leave a facility, he does not take with him all the knowledge and insight necessary in order to use the instrument effectively. This means that a standardized manual can help prevent "reinventing the wheel" or having to organize and develop the same work sample over and over. All these benefits point out that work sample standardization is a critical aspect of an evaluation program since it facilitates a systematic analysis of client

capabilities. Thus all clients are essentially given an equal chance to achieve maximum performance.

PART III*

Work Sample Norms

An individual's performance score on a work sample is most useful when other scores are available for comparison. The choice of a comparison score or group of scores should be guided by the underlying purpose for which a work sample is being used. Since work samples are commonly used as an aid in the assessment of a client's potential for functioning in various competitive occupations, it seems logical that work sample comparison scores based on performance rates of workers from competitive industry should be used whenever possible. Preferably, these competitive standards will reflect the production expectations for newly employed workers rather than experienced workers who in most cases perform at higher levels than inexperienced employees.

Part "c" of standard 3.4.3.1.1.7.2 states that when using work samples, "competitive norms or industrial standards shall be established and used" (CARF, 1978, p. 28). The CARF-VEWAA interpretation of this guideline further explains:

> There should be evidence (such as time studies, Methods-Time-Measurement (MTM), employer opinions as to relevance of work sample content or quality and quantity standards, actual competitive worker performance data, etc.) that performance standards used for interpretive purposes have been related to appropriate competitive performance standards. (Vocational Evaluation and Work Adjustment Association and Commission on Accreditation of Rehabilitation Facilities, 1978, p. 6)

The standard and the interpretation stress that all work samples used for assessing client potential for competitive occupations should have competitive norms or industrial standards available. In this case the terms norms and industrial standards both refer to the concept of an average (mean or median) performance score for a task or series of tasks as well as the range of scores deviating above and below the average. Competitive norms and industrial standards are extremely advantageous because they offer a comparison group for determining not only whether or not a client can perform a job, but <u>how well</u> the client can do the job. For example, if a client is given a simple nut and bolt assembly work sample without norms, the evaluator can only determine whether or not the client can actually correctly perform the assemblies, e.g., does he have adequate dexterity or can he remember the sequence of operations. The evaluator only has a limited idea of how well the client can do the job. The question as to whether or not he can perform the assemblies as fast as employed nut and bolt assemblers has still not been answered. This question can, however, be answered if the evaluator knows what the competitive production standards are. For example, if newly hired workers (e.g., less than 40 hours nut and bolt assembly experience) are expected to perform a minimum of 100 assemblies per hour, then the evaluator has a relatively easy job of comparing the client's performance to the competitive standards and thereby making more reliable decisions as to employment potential.

---

*Much of the material provided in Part III of this publication was taken with permission from the authors, C. Thomas Allen and Arnold Sax, of an earlier MDC publication entitled <u>Norms and Performance Standards for Work Sample Scores</u>.

Perhaps the most important point of this guideline is that competitive norms or industrial standards must be used. The term competitive denotes the concept that the norms or industrial standards should be based on people actually employed in occupations requiring the skills, traits, etc., required of the work sample rather than a handicapped population or general group of people for which there is no objective evidence as to whether or not they could really do the job(s). In addition to this, the industrial or competitive standard used in a work sample should reflect the standards used in industry for the same job. For example, an MTM study on a work sample might indicate that 300 pieces per hour is 100% competitive, and it can easily be erroneously assumed that this is the standard industry uses for worker performance. Yet, in actual practice, an industry might regularly employ people who perform at only 70% of the work sample's industrial standard. Thus the industrial standard used in the work sample should, as much as possible, reflect the actual practices used in industry.

Norms and Industrial Standards

Norms are needed in order to compare an individual's performance to other appropriate reference groups. With regard to vocational evaluation, such reference groups should generally be composed of workers employed in occupations closely related to the work sample. The usefulness of a norm group is determined to a great extent by the methods used to select its members. Care must be taken to insure that members of the norm group are representative of populations to which comparisons are to be made. For example, a work sample that assesses an individual's potential to work as a small engine mechanic would more appropriately have a norm group of successfully employed small engine mechanics or some other closely related occupational group, rather than a norm group of college students. Regardless of the degree of planning that goes into the development of norms, one must collect data on individuals that are available for testing. As a minimum requirement, one should provide an accurate description of members of the norm group. There are three different types of norm groups which generally apply to work sample testing:

1. Client norm groups
2. General population norm groups
3. Competitive norm groups or industrial standards

1. Client Norm Groups

In many work evaluation settings the most commonly used norms are those based on scores of other clients who have performed a specific work sample. These are not, however, competitive in nature and, therefore, do not satisfy CARF requirements.

Even though client norms may be relatively easily obtained, there are many serious problems with using them. Client norms are dependent upon the extent of abilities (or disabilities) of the clients served by a particular facility. Since client groups may vary considerably among particular facilities, norms based on clients from one facility may be misleading for use in another facility without detailed descriptions of clients included in the norm group. Also, within a particular facility norm groups may differ for each work sample. One work sample may be given to almost every client. Another may be administered only to those that an evaluator has judged from prior testing to function

at farily high levels. Thus, a client whose performance ranks relatively high compared to the norm group for one work sample may rank much lower as compared to the norm group of another work sample requiring essentially the same skills and aptitudes. Finally, the major problem with client norms is that a client's performance may appear to be above average when compared to other clients. Yet, if the client's performance was compared to competitively employed workers who represent the group the client will actually have to compete with for employment, his performance level might be considerably below competitive standards. Thus, using client norms can easily lead to the erroneous conclusion that a client has potential in a work area which may actually be inappropriate for him. Because of this problem, it is imperative that if an evaluator is considering the client for competitive placement, competitive norms must be used since client norms will not specify what performance standards are necessary for successful employment.

In spite of the limitations in using client norms, there is one instance where they can be very useful. That is when a client is being evaluated with regard to placing him in sheltered employment. In this case, the worker population which the client will be competing with is actually a client group, rather than competitive workers employed by industry. However, the workshop in which the client is to be placed must still have competitive production standards available for measuring their client's performance relative to the industrial standard. Thus, an evaluator might know that a sheltered employment program is involved in an assembly operation whose competitive industrial production standard is 100 pieces per hour. Workshop production records indicate that the average productivity of their clients, however, is only 40 pieces per hour or 40% of the industrial standard. So if the evaluator has a client whose performance is 50 pieces per hour on the work sample which was derived from the workshop task, the evaluator should recognize two important things. The client is producing below competitive industrial standards, but if placed in the workshop his production skills are compatible with the other sheltered employees. In this case, client norms would be useful for determining a new client's ability to produce at levels comparable to the standards required of a specific sheltered workshop.

2. General Population Norm Groups

The second type of norm group is based on a general population of non-disabled subjects. The evaluator administers a work sample to a random sample of non-disabled people who are not competitively employed in jobs related to the work sample. This group is referred to as a sample of the general population. In most cases, however, it is impractical for evaluators to obtain a true random sample of the general population. Instead, the non-disabled population is made up of a select group of individuals who are readily available, e.g., workshop staff, students, or volunteers. Like the use of client norms, this method of comparing client performance has serious limitations. Work sample norms based on the performance of members of a general, non-disabled population offer no direct basis for determining whether a client is capable of functioning at or near the rate of competitive workers on a particular work task, since none of these people are employed in a job directly related to the work sample, and it cannot be assumed that they would be successful workers simply because they are not disabled. Thus, predictions as to client performance capability are unsubstantiated and unreliable if they are based on either client or general population norms.

3. Competitive Norm Groups or Industrial Standards

As suggested previously, whenever an evaluator seeks a direct basis for determining potential for success in competitive situations, competitive norms or industrial standards based on the performance rates of workers in competitive industry should be used. These are the only techniques which satisfy CARF guidelines. This is because they provide a reliable and direct basis for comparison of client performance versus real work requirements.

Competitive norms are more difficult to obtain than noncompetitive norms; however, the benefits justify the effort. One method of obtaining competitive norms involves having workers employed in jobs closely related to the work sample actually perform the work sample. The results may then be statistically analyzed to obtain a relatively good picture of competitive performance requirements. For example, a group of competitively employed electronics assemblers would take an electronics assembly work sample which may not be exactly the same as their job but which involves the basic critical elements of the job. Based on the performance of this group, the evaluator should have a sound idea of how well competitively employed workers can perform on that particular task. This technique is particularly useful when an evaluator has a work sample which has already been developed, but lacks any normative data. It can also be useful for a work sample which is based on a job taken from industry, which has no concrete production standards. For example, a service occupation such as maid service may not have any specific production standards, but based on a job analysis, the evaluator develops a maid service work sample. In order to obtain norms for such a work sample, the evaluator may choose to have several people who are competitively employed as maids perform the work sample. Then the appropriate data such as speed, quality, etc., could be obtained and analyzed so as to provide performance criteria.

A second method of obtaining competitive norms or industrial standards involves developing a job sample. The job sample is probably the most effective and efficient way to obtain an industrial standard, because it is based on real work being done in industry. In this case, the evaluator contacts businesses and identifies jobs that hold potential for the client population being served. A job analysis is performed, and then the evaluator can create a job sample which replicates the real job in its entirety. Once the job is accurately simulated, the evaluator may then use the same production standards being used by the industry from which the job was taken. In effect, the evaluator simply borrows a competitive production standard which industry has taken the time and trouble to develop. This approach can save the evaluator a great deal of time and effort since the problems of soliciting volunteer workers from industry, accumulating data, and synthesizing/analyzing results are eliminated. The second advantage is that rapport is established with potential client employers and specific jobs are pinpointed. The evaluation unit may, therefore, serve as a more effective placement tool. Third, the job sample inherently has a great deal of validity since it fully replicates a job, with the possible exception of matching the entire industrial environment. Fourth, industry should be able to specify such factors that might be observed in the first few days which may distinguish between those new employees that eventually become good producers from those who are marginal or unacceptable in production. It should be emphasized, however, that it is imperative that if production standards are taken from industry, the job sample must replicate the job in every way. The same production methods, layouts, tools, equipment, etc., must be used for both the job sample and

the job. Otherwise, the standards may no longer be valid since different procedures may strongly influence performance (for a further discussion of the job sample, refer to Piller, 1977).

The third method of obtaining competitive performance standards for work samples involves establishing production standards which directly apply to the work sample. If an evaluator has a work sample which needs to be normed, but neither of the first two techniques are possible, e.g., (1) there are too few competitively employed workers available for taking the work sample or the jobs are too dissimilar, or (2) the production standards used by industry can't be applied because a job sample couldn't be developed or the industry had no specific production standards, then production standards may still be established by using work measurement techniques borrowed from industry (the different kinds of work measurement techniques will be discussed in the following section). Rather than bringing workers in to take the work sample, or developing job samples, this procedure requires that the evaluator contact a qualified engineer to perform time studies, Methods-Time-Measurement, Master-Standard-Data, or other work measurement techniques to arrive at reliable, objective production standards for the work sample (refer to Botterbusch, 1975, for further discussion). Although this procedure is useful, it also has a significant limitation in that it may be very costly to have an engineer come in and analyze a work sample and develop production standards. This is particularly true in cases where many work samples are involved or individual work samples are very lengthy operations of several minutes or even hours. It may take several hours just to analyze a two or three minute job, and this problem must be taken into consideration when selecting a method for obtaining performance standards. One way to minimize this expense is to try and obtain qualified volunteers through community involvement or by contacting appropriate professional organizations. Interested parties may wish to get in touch with the:

American Institute of Industrial Engineers, Inc.
25 Technology Park/Atlanta
Norcross, Georgia 30092

They may be able to provide further information with regard to locating the nearest chapter.

Work Measurement Applied to Work Samples

Industrial standards are usually derived from a variety of work measurement techniques. Work measurement techniques have been developed by industrial engineers and used successfully to estimate labor costs for future projects, schedule production, evaluate different methods of performing a job, and establish incentive wages. Work measurement involves the detailed analysis of a particular task using standardized techniques in order to determine the amount of time needed for an <u>average or normal experienced</u> worker to perform the task under normal conditions. The industrial engineer is concerned not only with establishing an accurate estimate of the time needed to perform a task, but also in evaluating the most effective methods used to perform the task. Methods of setting up the work task are examined to eliminate such factors as awkward motions, long reaches, and excessive material handling (for a film review of this subject, refer to Botterbusch, 1975).

The two most popular systems of work measurement are the time study and predetermined motion time systems. Each type offers specific advantages depending on a variety of circumstances and the types of work tasks to be studied. Time study is a method of determining how much time a task should require. A clock, stopwatch, or other timing device is used, and the time that elapses while the task is being done is recorded. Usually a series of these observations or timings take place using a group of average, experienced workers. An average time is then determined from the performance of individuals within the reference group, and this time is thought to represent the normal amount of time required for an average experienced worker to perform the desired operation(s). A predetermined time system is an industrial engineering technique that allows production time standards to be determined without relying on an average, experienced worker. Instead of timing the motions of individual workers, the specific motions required to do the job are identified and listed on an analysis sheet. A time is then assigned to each of these movements. These times are taken from standardized tables whose values have been predetermined, i.e., thousands of manual motions were observed and broken down into a few well-defined motions whose performance time was statistically analyzed and averaged. The theory behind these systems is that elements of human movement, on the average, take the same amount of time to perform, regardless of the individual worker involved. These systems are quite reliable and are widely used by government and industry. For further discussion of this subject, refer to Botterbusch (1975).

Because time study requires the observation of average, experienced workers which are generally not available in most workshop or evaluation settings, predetermined motion time systems are often more appropriate for developing performance standards for work samples. The exception to this is the case where the evaluator can bring in average, experienced workers from industry or workshops and use them as the basis for the time studies. However, when average, experienced workers are available, corresponding production standards will usually already have been developed since the job is already in progress. Thus, the evaluator may be able to save time and money by simply borrowing the standards already developed by industry. Of the several standardized predetermined motion time systems, Methods-Time-Measurement (MTM) appears to be the most widely known and commonly used. Some other systems include Modular Arrangement of Predetermined Time Standards (MODAPTS) and Master-Standard-Data. For more information the reader should contact the nearest industrial engineering organization.

Work measurement may be used to provide valid and realistic performance standards for many work samples, especially those involving repetitive manual operations. The work sample need not be a work task performed in industry. Work measurement may be used regardless of whether a work sample is a simulation of a specific industrial job or a task which is intended to relate to a variety of jobs. If the work sample task were to become an actual job in industry, the work measurement results should be a good estimate of the performance rate of an "average" worker. Since there is no immediate need to gather normative data, new work samples for which work measurement standards have been developed are immediately usable for evaluation purposes.

A client does not have to meet or exceed the industrial performance standard in order to be considered to have a good chance for success in performing tasks required in jobs related to a work sample task. Performance equal to the standard is designated 100% performance. If the task is a type in which the establishment of work rhythm is important, performance at 30% to

50% of standard might be considered quite good on the first trial of the work sample. Due to varying levels of physical fatigue and emotional stress in the evaluation situation, a client's work performance may differ considerably from time to time. It is generally preferable that a client be administered the work sample several times. Since a work measurement standard is based upon the performance of the average experienced worker, repeated trials of a work sample are generally advised. Repeated trials allow opportunity to observe signs of interest in the work tasks. Close supervision is generally not required after the client has learned to perform the task correctly. It would be dangerous to specify a percentage of an average competitive rate that might be used as an absolute lower limit for judging that a client has potential for successful employment. Any guidelines used must be tempered by consideration of variables such as task complexity, amount of practice, and client anxiety. Beginning and marginal workers generally perform at rates of at least 70% of standard. Although this 70% figure might be used as a rough guideline, it should not be considered a firm cutoff point. An industrial engineer who develops a performance standard should be able to provide general guidelines in light of complexity of the task and local industrial expectations.

Cautions in Using Competitive Norms or Industrial Standards

When using competitive norms or industrial standards for interpretive purposes, there are certain cautions which must be observed. First, if an evaluator administers a work sample to a competitively employed group of workers, he must be certain that the work sample is very closely related to the job performed by members of the reference group, i.e., an electronics soldering work sample should be administered to electronics solderers and not a group of heliarc welders. Otherwise, the evaluator might find that he has competitive norms but they are essentially useless for predicting job potential. For example, if an evaluator has a cashiering work sample, he should not select several clerk typists for the norm group. Such a group would be inappropriate since no matter where a client's performance was in comparison to the competitively employed norm group, there would be no evidence to suggest that the client could or could not work competitively as a cashier since there is no evidence that the clerk typists would be successful cashiers. In such a case, all that could be said is that in comparison to a group of clerk typists, the client's performance score fell at a certain level. Although, in this case, competitive norms have been used, the information derived from these norms is useless and could be misleading.

Second, when work measurement is involved, one must not only be concerned with the time needed to perform a task, but also the methods used to perform the task. In order for work measurement standards to remain valid, it is imperative that the method used while establishing the industrial standard be precisely duplicated by the work sample method. This means all the motions, distances, tools, etc., must be exactly the same. If the method is changed, the previous standard must be replaced by an updated standard based on the new method. Regardless of the type of instructions used for the work sample, the client must demonstrate his competency in correctly performing all components of the task in the standardized manner before his speed of production is assessed (McCray, 1979). Instructions or demonstrations should be repeated or modified if necessary to insure that the client performs the task correctly.

Third, once a competitive standard is obtained for a work sample, the evaluator must not be content with accepting it alone as a rigid, inflexible criteria for predicting successful employment. Instead, the evaluator must be aware of what the prevailing production standards are for competitively employed workers in similar jobs in the surrounding community, since the actual standards used by industry may be different than the standards for the work sample. For example, although an MTM study might indicate that 100% competitive on an assembly work sample is defined as 50 pieces per hour, industry may be willing to employ workers who are substantially less than 100% productive. Successful employees may be producing at only 70% of the standard or 35 pieces per hour. Or a business may hire new employees who are only 50% productive during the first few weeks of training. Thus, as much as possible, the evaluator must be aware of the real performance requirements of industry if a reliable prediction as to employment potential is to be obtained.

Finally, a major area of concern in using industrial standards is that these standards are based on average, experienced workers. In some cases this may mean that the average competitively employed worker is allowed hundreds or even thousands of practice trials before they are expected to achieve a minimum performance standard. Thus, comparing an inexperienced client's performance to the standards of experienced workers, and using this as the basis for predicting vocational potential can be misleading. Under these conditions, one can only be relatively certain of an individual's performance potential when he scores high in comparison to the industrial standard. For example, if on the first few trials a client achieves 80% of the industrial standard expected of an experienced worker, the evaluator can be reasonably sure that the client has the skills and aptitudes required to do the job. Yet, if a client only scores 50% of the industrial standard, the evaluator cannot be certain that the client does not have the potential to do the job, and it is unrealistic to provide the client with an indefinite number of additional trials. Unless the evaluator knows how many trials it typically takes to achieve competitive performance, the evaluator cannot assume that a client's potential in a given area is limited. For example, even though 50% of the industrial standard may seem low, a closer look at the standard might indicate that it usually takes 100 trials for inexperienced workers to achieve this level of proficiency. Perhaps the best way to resolve this dilemma is to obtain local industrial standards for inexperienced or new employees as well as experienced workers. Thus, a local industry may regularly hire and employ inexperienced workers whose production standard during the first 200 hours of work is only 50% of the experienced workers' standards. Such an approach allows the evaluator to compare an inexperienced worker's performance with standards for inexperienced as well as average experienced workers.

When used by adequately trained and experienced industrial and vocational specialists, work measurement techniques are both reliable and valid. Although errors of judgment can and do occur on occasion, well trained and experienced work measurement specialists are generally able to estimate production standards with error rates of 5% or less. Labor unions have generally accepted work measurement performance standards, when agreed safeguards are provided to assure accuracy and there are opportunities to request reexamination of jobs for which standards appear to be in error. Thus, it is evident that work measurement techniques can provide relatively objective and accurate methods for determining competitive work sample performance standards.

SUMMARY

A work sample can be an invaluable asset in assessing the vocational skills, interests, aptitudes, behaviors, etc., of handicapped people. However, this is only true when it is constructed in such a way as to be a representative, systematic, and reliable simulation of real work. VEWAA and CARF, by establishing some concrete guidelines directly related to work sample testing, have taken a step toward assuring that these essential characteristics are incorporated in all work samples.

When evaluating the quality and utility of an individual work sample or a work sample battery, evaluators should examine them with a critical eye. Poorly developed work samples can have a profoundly negative impact on the effectiveness of a vocational evaluation program. At the very least, a work sample should:

(a) be representative of realistic competitive worker skills
(b) be based on job analysis techniques
(c) be standardized
(d) have appropriate competitive or industrial norms.

If a work sample is able to meet the aforementioned criteria, it may be regarded as an important asset in assessing vocational potential. Thus, evaluators, clients, and referral sources will have one more reliable technique for gaining an accurate and meaningful understanding of client capabilities.

REFERENCES

Allen, C., & Sax, A., Norms and performance standards for work sample scores. Menomonie, Wisconsin: University of Wisconsin-Stout, Department of Rehabilitation and Manpower Services, Materials Development Center, 1972.

Appalachia Educational Laboratory, Worker trait group guide. Bloomington, Illinois: McKnight Publishing Company, 1978.

Botterbusch, K. (Producer). Industrial Engineering in Rehabilitation. Menomonie, Wisconsin: Materials Development Center, 1975. (Film)

Commission on Accreditation of Rehabilitation Facilities, Standards manual for rehabilitation facilities. Chicago: Author, 1978.

Dunn, D., Validating work evaluation procedures. Paper presented at Short-Term Training Program, University of Wisconsin-Stout, Menomonie, Wisconsin, 1971.

Manning, W. H., Test validation and EEOC requirements: Where we stand. Personnel, 1978, 55(3), 70-77.

Materials Development Center, Work sample manual format. Menomonie, Wisconsin: University of Wisconsin-Stout, Stout Vocational Rehabilitation Institute, 1977.

McCray, P., Learning assessment in vocational evaluation. Menomonie, Wisconsin: University of Wisconsin-Stout, Stout Vocational Rehabilitation Institute, Materials Development Center, 1979.

Piller, R. (Producer). Introduction to the Job Sample. Menomonie, Wisconsin: Materials Development Center, 1977. (Slide-Sound)

U.S. Department of Labor, The dictionary of occupational titles (Third Edition) (Two Volumes). Washington, D.C.: GPO, 1965.

U.S. Department of Labor, The dictionary of occupational titles (Fourth Edition). Washington, D.C.: GPO, 1977.

Vocational Evaluation and Work Adjustment Association and Commission on Accreditation of Rehabilitation Facilities, VEWAA-CARF Vocational evaluation and work adjustment standards with interpretive guidelines and VEWAA glossary. Menomonie, Wisconsin: University of Wisconsin-Stout, Stout Vocational Rehabilitation Institute, Materials Development Center, 1978.